International Journal of Computational Intelligence in Control

Detecting Rising Stars: Evidence from Top Pakistani Universities by using Co-author, Power graph and Datamining Techniques

Mamoona Anam
Faculty of Computing and Information Technology, Department of Computer Science
International Islamic University, Islamabad Pakistan 192122 mamoona.anam.vt@iiu.edu.pk

Malik Sikander Hayat Khayal
Department of Computer Science, Preston University, Islamabad Pakistan 192122
drsikandarhayat@preston.edu.pk

Bashir Khan
PhD Scholar
b.khan@hec.gov.pk

Muhammad Asghar
Faculty of Management Sciences, International Islamic University, Islamabad Pakistan 192122
muhammad.asghar@iiu.edu.pk

Received: 28th September 2021
Revised: 20th October 2021
Accepted: 19th November 2021

***Abstract:** The process of assessing the scholarly output of academics is becoming increasingly challenging within the contemporary landscape of academia. Evaluation committees often extensively search multiple repositories to compile their evaluation summary report on an individual. Nevertheless, deriving performance metrics about a scholar's dynamics and progression poses a considerable challenge. This study introduces a novel computational approach utilizing unsupervised machine learning, which has the potential to serve as a valuable tool for committees tasked with evaluating the scholarly achievements of individuals across different universities of Pakistan, namely, Air, Quaid e Azam, International Islamic University, FAST, UET Taxila, COMSAT and NUST university. The proposed methodology generates a comprehensive set of key performance indicators (KPIs) for each researcher and monitors their progression over time. The considered variables are employed within a clustering framework, which uses clustering validity metrics to automatically ascertain the optimal number of clusters. This is done before the classification of scholars into distinct groups. The assignment of performance indicators to the clusters can ultimately function as the primary profile characteristics of the individuals within those clusters. This enables the deduction of a profile for each scholar. The present empirical investigation centres on analyzing rising or emerging stars who exhibit the greatest advancements over time concerning all Key KPIs. Additionally, this study can be utilized to assess the performance of scholarly groups.*

***Keywords:** Rising stars, Power graphs, Co-authorship graphs, Data mining, Key performance indicators, Pakistan*

1. Introduction

Evaluating the performance of research scholars and faculties is becoming so difficult in today's modern era of academia(Fitzgerald, Karen, Sonka, Furco, & Swanson, 2020).The main reason for this difficulty is the different objectives or criteria various departments set for each faculty. There has been a notable transition in the composition of faculty members within colleges and universities worldwide in recent years. Specifically, there has been a shift away from the traditional model of employing full-time, tenure-eligible faculty members that was prevalent in the past. Instead, there has been a growing reliance on part-time faculty members, non-tenure-track faculty members, and instructors who primarily teach online courses. Furthermore, there exists a dearth of universally acknowledged standards in this domain. Instead, there are many recommended approaches, frequently disseminated by individual academic departments, regarding evaluations(Panagopoulos, Tsatsaronis, & Varlamis, 2017).

Assessing the research productivity of faculty members and scholars presents a multifaceted undertaking accompanied by a range of inherent difficulties. Several prominent challenges can be identified. The first is the subjectivity. The research evaluation process entails subjective assessments, as evaluators may possess diverse criteria and interpretations regarding the quality of research. The inherent subjectivity of the evaluation process can potentially give rise to biases and inconsistencies. The second is the multidimensionality. It refers to the diverse aspects that comprise research performance, including but not limited to publications, citations, grants, collaborations, and societal impact. Formulating comprehensive evaluation metrics that effectively encompass the multifaceted aspects of research can present difficulties. The time and lag effects in the study reveal that the research process is a laborious and protracted endeavor, often spanning several years before the consequential influence of research outputs becomes discernible. Assessing the performance of researchers in the immediate term may not yield a comprehensive depiction of their enduring scholarly impact(McKenney & Reeves, 2021).

Further, it is worth noting that there may exist a temporal delay between the dissemination of research findings and their subsequent citation or measurement of results. The fourth challenge is that field-specific variations refer to the unique research practices, publication norms, and impact metrics within different academic disciplines. The assessment of researchers across various disciplines necessitates a comprehensive comprehension of these disparities and the capacity to accommodate them appropriately. Data availability and reliability are the fifth factors that pose a significant challenge when accessing accurate and comprehensive information to evaluate research performance. Various data sources may possess inherent limitations, and it is not always feasible to quantify or measure all research outcomes using conventional metrics. Further, interdisciplinary research has increased; however, existing evaluation systems are primarily tailored for assessing single-discipline endeavours. The assessment of the impact and contributions of multidisciplinary research poses challenges due to the absence of well-defined evaluation frameworks(McKenney & Reeves, 2021).

Furthermore, unintended consequences in gaming evaluation systems are a subject of concern. These systems have the potential to inadvertently incentivize researchers to prioritize quantity of output rather than quality, exhibit bias towards specific types of publications, or even engage in unethical behaviors. Evaluators must exercise caution regarding these potential challenges and make diligent efforts to establish equitable and rigorous evaluation methodologies(Marchionini, Plaisant, & Komlodi, 2003).

Moreover, the evaluation methods commonly employed in academia may not sufficiently encompass the varied contributions made by researchers from underrepresented groups or those engaged in non-conventional research outputs, posing limitations to assessing diversity and inclusivity. Promoting equity, diversity, and inclusivity in the research assessment necessitates meticulous deliberation regarding the criteria and metrics employed in the evaluation process. To effectively tackle these challenges, employing a comprehensive approach that encompasses meticulous planning, open communication, active involvement of relevant parties, and ongoing enhancement of assessment methodologies is imperative. Enhancing evaluation frameworks that are more comprehensive and nuanced can assist in addressing these challenges and offering a more extensive comprehension of research performance. Consequently, the evaluation committees frequently have to generate evaluation summary reports for individuals consistent with the faculty's scope and objectives, despite lacking the opportunity or time to conduct an in-person meeting with the candidate. To fulfil this objective, there is a significant requirement for automated tools that can evaluate scholars' profiles, encompassing a diverse array of Key Performance Indicators (KPIs) that address the majority of the faculty's objectives and requirements(Bornmann & Marx, 2015; Hicks, Wouters, Waltman, De Rijcke, & Rafols, 2015; Waltman & van Eck, 2013).

Many KPIs can be considered when assessing the performance of academics and faculties. The evaluation may encompass various aspects, including teaching and research performance, capacity to secure research funding, and the acquisition of patents, whether awarded or pending. This study centers on the assessment of individual scholars based on their performance in research activities. In our argument, we contend that when evaluating the scientific performance of individuals, it is crucial to consider bibliometric indicators such as the h-index(Hirsch, 2005) and journal impact factors, as well as the social aspects of the researcher's contribution and its dynamics. The argument presented in this paper is supported by previous research that has identified inconsistencies in the h-index(Waltman & Van Eck, 2012) and impact factor(Seglen, 1997). Additionally, it is argued that using a single-number evaluation model has limitations, such as restricting the consideration of alternative interpretations and failing to adequately highlight a scholar's strengths, weaknesses, and future potential. This perspective would also facilitate the evaluation committees in thoroughly documenting the underlying reasoning behind their decision and aligning it more seamlessly with the objectives of the faculty.

To fulfil the criteria above, we propose a computational approach for examining the profiles of individual scholars. This approach analyzes performance indicators encompassing scientific achievements and social/collaborative aspects. Additionally, we investigate the evolution of these indicators over time. Various metrics are utilized to assess an author's productivity and the impact of their work, including the number of publications and the number of citations received by each publication. These metrics are particularly focused on the transient nature of these measures, allowing for a distinction between works that have a lasting influence and those that achieve temporary success. In addition, assessing an author's sociability involves the examination of their co-authorships and the influence these collaborations have on their professional trajectory. This analysis is conducted through collaboration networks, specifically co-authorship graphs, constructed from bibliographic data. These graphs are then subjected to traditional graph mining techniques(Cook & Holder, 2006) and Power Graph Analysis (Royer, Reimann, Andreopoulos, & Schroeder, 2008). Subsequently, we quantify the variations of each indicator over successive time intervals to capture the dynamics of the scholarly field. In the final stage, the scholars' profiles are clustered within each period by utilizing the indicators mentioned earlier as features. Subsequently, a feature analysis is conducted to delineate the characteristics of the clusters, followed by an evaluation utilizing the future values of the clusters.

The ethical implications associated with the assessment of academic performance have been extensively examined and discussed in the existing scholarly literature. As an illustration, Sir Philip Campbell, the Editor-in-Chief of the esteemed Nature journal, asserted that the most efficient and equitable evaluation of an individual's contribution is best derived from a direct assessment of their papers rather than the specific journals they were published in(Campbell et al., 2010). The phenomenon above resulted in a decrease in the significance attributed to impact factors and an increase in the importance placed on individual citations per publication (Cronin & Overfelt, 1994). The h-index, a well-known metric, has faced criticism due to its potential to present misleading information about a scientist's productivity and susceptibility to various biases. These biases include the inclusion of self-citations, the publication of research in different domains, the influence of open-access publishing, and the cumulative advantage experienced by more senior researchers. Using extensive bibliographic and citation databases, such as Scopus and Google Scholar, which offer precise citation counts for each publication, has effectively addressed the biases related to venue impact and the cumulative effect of older papers. However, certain challenges remain unresolved, such as accurately quantifying the individual contributions of authors in a publication.Retzer and Jurasinski (2009)examine the challenges associated with research evaluation and propose a strategy to enhance objectivity in evaluating evaluation metrics. Another aspect currently being extensively studied is assessing the true quality of scientific work. This complex and multifaceted task goes beyond mere consideration of publication and citation counts(Senanayake, Piraveenan, & Zomaya, 2015).This article presents a proposed set of indexes for assessing an author's research potential. The methodology involves adjusting the impact of work over time and analyzing the temporal changes in these indexes. This approach mitigates the cumulative bias inherent in traditional ranking or classification methods. Instead, the authors advocate for clustering authors with similar potential into distinct groups. This categorization can enhance the efficiency of the researcher selection process by narrowing down the pool of candidates to a smaller subset of individuals who demonstrate strong potential. This allows for a more comprehensive selection process that considers various aspects of each candidate rather than solely relying on quantitative metrics and indices.

Our analysis focuses on the cases of emerging talents, given their significance and the complex challenges they present. According toMa et al. (2020)rising stars refer to individuals who currently possess relatively modest profiles but have the potential to become prominent researchers in the future. The challenge in identifying emerging talents involves making predictions over multiple years instead of an immediate classification task. To make an early assertion regarding the future success of an author as a prominent scientist, it is imperative to gather data during a period when the author is relatively unknown. Therefore, it is necessary to examine specific details that reflect the potential of this author rather than relying on traditional metrics for evaluating authors.

Moreover, it is crucial to monitor the author's publication and citation history over time to verify the accuracy of the prediction. In a prior study conducted byPanagopoulos et al. (2017), an endeavor was made to construct models for the profiles of group leaders. Additionally, the concept of rising stars was introduced, referring to authors who exhibit significant growth in both the quantity and influence of their published work as time progresses. According toDaud et al. (2020) rising stars are defined as authors who can potentially become prominent contributors in the future, despite currently having a low research profile.

The primary contribution of this study is the integration of quantity, impact, and collaboration-related features, which have not been combined and monitored over time in previous research. This approach allows for creating scholar profiles that effectively identify their strengths and weaknesses. Authors are categorized into subsets of relevant performance based on a clustering algorithm, cluster validity measures, and a set of time-evolving features. The clusters are assigned labels based on their most representative features. The features with the highest values across clusters are utilized for each cluster. One further contribution of this study is the employed methodology for cluster analysis, which effectively emphasizes the specific topics of interest to the authors within each cluster, the publication forums they typically utilize, and the affiliated domains.

The analysis conducted has yielded three significant observations. First and foremost, it can be observed that scholars who share similar values in terms of bibliometric Key Performance Indicators (KPIs) also exhibit similar values in their collaboration KPIs. Furthermore, it is imperative to closely observe the progression of key performance indicators (KPIs) over time to discern scholars who have consistently generated influential works and demonstrated enduring significance. Finally, within the identified clusters, we could discern groups of scholars who demonstrate significant levels of dynamism, indicating their potential to achieve elevated levels of success in their respective careers. These academics are frequently called "emerging talents" in the scholarly discourse.

2. Literature Review

Recently, numerous studies have aimed to identify emerging scholars within academic networks. The literature in question employs various methodologies to evaluate authors, including assigning scores to gauge their potential for success, categorizing them intorising or non-rising classifications, or grouping authors with comparable attributes into clusters. Most studies extract multiple features or indicators for each author, which are subsequently assessed for their efficacy in distinguishing individuals with potential for future success. However, thus far, none of the studies have utilized a comprehensive approach incorporating all three elements, specifically bibliometric indicators, collaboration indicators, and longitudinal monitoring.

To identify emerging stars more specifically, offer the Pub Rank method. Research on the evaluation of researchers and prediction of rising stars is getting attention because it can be useful for selecting capable candidates for the jobs, hiring young faculty members for institutes, and seeking reviewers for journals and conferences and members for different committees.Urooj, Khan, Iqbal, and Alghobiri (2023) address the research problem of finding rising stars and propose novel features in diverse feature sets of three categories: article, author, and venue. The real-world data set has been extracted, preprocessed, and used from the Web of Science for empirical analysis. Several diverse supervised machine learning, ensemble learning algorithms, and deep learning are applied to the data set.

There are two main issues to address in this context. Firstly, the author's contribution to a publication is being ignored. Secondly, viewing the venue rank as a static measureDaud et al. (2020) is inappropriate. Identifying emerging talents, commonly called Finding Rising Stars (FRS), has become a prominent subject of investigation in various fields of study. In contemporary times, there is a prevailing inclination among individuals to prioritize identifying individuals who possess the potential to acquire the expertise to occupy junior-level positions rapidly rather than focusing on recruiting established experts who can promptly assume senior-level roles. The presence of FRS individuals can enhance productivity in any setting due to their dynamic and enthusiastic demeanour. This paper aims to evaluate the methodologies employed in the identification of FRS. The current methodologies can be categorized into ranking-, prediction-, clustering-, and analysis-based approaches, and an examination of the advantages and disadvantages of these approaches is presented. This growing area of research also includes comprehensive information on standard datasets and performance evaluation measures. In conclusion, we will now address the remaining challenges and potential avenues for future exploration within this thriving field of study(Daud et al., 2020).

Panagopoulos et al. (2017) introduced the Star Rank algorithm to solve the earlier limitations. This algorithm addresses these limitations by considering the order of author names and by dynamically adjusting the rank of venues over time based on the entropy of topics discussed in those venues. The first hypothesis is not universally applicable across all disciplines. For instance, in certain cases, co-authors may arrange their names alphabetically, while in other instances, the senior author is consistently positioned at the end of the list of authors. The second hypothesis is deemed weak as it fails to consider that the terms used in paper titles may not always accurately represent the breadth of topics covered within a particular academic venue.

Furthermore, the algorithm applies an equitable approach to conference and journal publications. Publishing in a journal with a significant impact is potentially more challenging compared to securing a spot in a prestigious conference. In their study, Long, Lee, and Jaffar (1999)employ a venue ranking scheme that exclusively considers conference rankings. However, whether journal impact factors or another metric is utilized for ranking journals remains unclear. Regardless of the circumstances, academic journals should carry more weight than conferences regarding an author's career. The experimental assessment resembles the study conducted byLi, Foo, Tew, and Ng (2009)as Long et al. (1999)manually evaluate the top ten predicted emerging talents at a future time point by considering the number of publications and citations. This evaluation serves as evidence to demonstrate the superiority of their approach over the Pub Rank mentioned above.

Al-Hoorie and Vitta (2019)This report analyzes the statistical methodologies employed by 30 journals that represent the second language field. An analysis of 150 scholarly articles revealed several common statistical violations, such as inadequate reliability reporting, validity, non-significant findings, effect sizes, and assumption checks. Additionally, researchers were found to draw inferences from descriptive statistics and neglect to address the issue of multiple comparisons. The predictive factors for the statistical quality of a journal include the Scopus citation analysis metrics and the journal's inclusion in the Social Sciences Citation Index (SSCI). Insufficient empirical support established a preference for the recently introduced Cite Score metric over SNIP or SJR. The discussion focuses on the implications of the obtained results.

Gibson, Anderson, and Tressler (2014) The assessment of an academic journal's ranking holds significant importance for authors, universities, journal publishers, and research funders. The significance of rankings has increased as nations increasingly implement periodic research evaluation initiatives that incentivize publication in prestigious journals with high-impact factors. However, even within a field that places significant emphasis on rankings, such as economics, there is a lack of consensus regarding the extent to which lower-ranked journals should be given less weight and the breadth of the journal universe that should be considered. Furthermore, authors may find it more economically advantageous to include unnecessary references, either voluntarily or under the influence of editors, rather than disregarding pertinent ones. Consequently, rankings that rely on citations can be susceptible to manipulation.

On the other hand, when considering the advantages of publishing in a particular journal versus another, evaluators are engaged in a deliberation process that carries significant implications for hiring, promotion, and salary determinations. Hence, our focus is directed toward the academic labor market, wherein we examine the correlation between economists' lifetime publications in 700 distinct academic journals and their corresponding salaries, utilizing data from economists within the University of California system. We analyze journal rankings and publication discount rates to determine which aligns most closely with the returns the academic labor market indicates. The user has provided three reference codes: JEL A14, I23, and J44.

Fu, Song, and Chiu (2014) present novel dimensions of an author's profile, specifically Influence, Connections, and Exposure. These dimensions offer alternative methods for ranking authors and provide a more comprehensive understanding of authors when combined with Citation Count. The research expands upon an intricate network structure that links various nodes (such as authors, venues, and papers) through different edges (including directed and undirected, bi-partite, and uni-partite connections). Additionally, it offers distinct author rankings for each metric. In their evaluation, the researchers examine the relationship between various metrics and their correlation with the h-Index. This study also fails to account for the temporal evolution of metrics, limiting its ability to identify authors with significant potential while still effectively identifying established authors.

Tagarelli and Interdonato (2015)conducted a study that focused on identifying inactive and non-productive users in social networks through time-aware analysis. The individuals commonly referred to as "lurkers" or "silent users" are characterized by a collection of productivity and content consumption attributes considering time. The authors also conduct clustering analysis on the entire population of users to identify the cluster of individuals who do not actively engage or participate.

The technique we suggest varies in numerous ways from the methods above:

1. According to Panagopoulos et al. (2017),assessing a paper's influence is determined by the number of citations it garners rather than the impact factor of the publication venue in which it is published. A paper's visibility and citation count can be significantly influenced by the venue in which it is presented. However, Seglen (1997)argues that the utilization of impact factors masks the variation in citation rates among articles. Based on the study above, it has been observed that articles belonging to the top 50% of citations in a journal receive citations at a rate that is tenfold higher compared to articles in the bottom 50% of citations. Consequently, the utilization of citations proves to be a more efficient and detailed metric for assessing the influence of a research paper in comparison to the impact factor of the publication venue.
2. Furthermore, in this study, we introduce the concept of a decay factor, which penalizes the level of success of publications by considering both their age and the age of the citations they have received. This adjustment enhances the robustness of our methodology in mitigating the self-citation bias while still allowing authors to cite their work for communication. Based on the prior research conducted by Costas, van Leeuwen, and Bordons (2010), it has been observed that self-citations tend to age faster than foreign citations. Consequently, the impact of self-citations on the cumulative influence of a publication diminishes rapidly, as self-citations typically emerge within a few years following the publication. The methodology presented in this study provides a more comprehensive understanding of an author's dynamics compared to the h-index proposed byHirsch (2005). Unlike the h-index, which solely considers the cumulative number of citations, our methodology considers an author's annual citation change. Therefore, it provides a more accurate differentiation among authors who consistently receive the same number of citations each year, authors who experience a steady increase in annual citations as they gain prominence, and authors who, despite a decrease in citations over time, possess a substantial cumulative citation count and an h-index exceeding 30 (Costas et al., 2010). Using clustering enables our methodology to form clusters of authors who share similar characteristics initially, then assign labels to these clusters and identify the one that consists of promising individuals.
3. In addition, we determine the most suitable number of clusters and conduct clustering analysis on the authors' dataset rather than binary classification. This approach enables

us to uncover additional categories of researchers, similar to the methodology employed byPanagopoulos et al. (2017).

4. In addition, we suggest a more rigorous qualitative assessment compared to the evaluation conducted byLi et al. (2009). This assessment will investigate the persistence of identified clusters over time and the extent to which these clusters maintain their overall levels in the most critical characteristics.
   The data utilized for this purpose is derived from the same dataset across multiple years. Specifically, the dataset is divided into two distinct portions: one for analysis and clustering and the other for evaluating the effectiveness of our algorithm.

3. Data and Methodology

This section is divided into two sub-sections. The first sub-section named Data describe the details about the data and sample collections. The second sub-section named Methodology explains the applied methods, techniques and graphical methods for this analysis.

3.1. Data

The dataset was generated by collecting data from the digital library Scopus, encompassing information about authors, their works, publications, and citations. One advantage of Scopus is its ability to provide an annual count of citations received by a publication instead of relying on impact indicators such as journal or conference impact factors. We considered that the authors in our research group were associated with Pakistani universities, specifically Air, COMSAT, Quaid e Azam, International Islamic University, FAST, UET Taxila, and NUST University.

3.2. Methodology

This paper contains exclusively focus on finite groups. The vertex set of the undirected power graph of a finite group G consists of the elements of G. Two distinct elements in this graph are considered adjacent if one element is a power of the other.The concept of a power graph was introduced byKelarev and Quinn (2000). The concept of an undirected power graph was introduced byChakrabarty, Ghosh, and Sen (2009). Numerous intriguing findings about power graphs have recently been acquired, as evidenced by the references.

Ayesha (2020) proposed the concept of the intersection power graph for a finite group. In his work, he introduced the concept of the intersection power graph.$\Gamma_{IP}(G)$ for each finite group called G.This graph is constructed by considering the group elements as the vertices and two distinct vertices a and b are connected in $\Gamma_{IP}(G)$ if $\langle a\rangle \cap \langle b\rangle \neq \{e\}$ and e is adjacent to all other vertices of $\Gamma_{IP}(G)$, where e is the identity element of G.

In this discourse, we shall revisit certain terminologies pertaining to graphs. A graph containing an edge connecting every pair of distinct vertices is called a complete graph. For the entire graphical analysis with n vertices, we use the notation.K_n. An independent set of a graph $\mathcal{G}$ is defined as a subset A of its vertices such that the induced subgraph on A contains no edges.

The independence number of a graph $\mathcal{G}$, denoted as $\alpha_0(\mathcal{G})$., is defined as the maximum cardinality of an independent set in $\mathcal{G}$. The notation $\alpha_0(\mathcal{G})$ will be used to represent it. A graph that is connected and contains only a single cycle is referred to as a unicyclic graph. A graph that lacks any edges is referred to as a null graph. The diameter of a graph is defined as the maximum distance between any two vertices within the graph. A friendship graph, denoted as F_n, is an undirected planar graph with $2n + 1$ vertices and $3n$ edges. The friendship graph can be constructed by connecting n instances of the cycle graph.C_3 at a shared vertex.

A book can be defined as a compilation of half-planes, each sharing a common boundary line. Arranging a graph in a planar manner within the pages of a book is referred to as book embedding. The book thickness of a graph $\mathcal{G}$ refers to the minimum number of half-planes required for any book embedding of the graph. The notation $bt(\mathcal{G})$ will be used to represent it.This paper aims to investigate issues related to unicyclicity and establish certain bounds for the diameter, book thickness, and independent number of the intersection power graph in specific scenarios. The motivation for these findings is derived from previous studies, specifically reference(Aalipour, Akbari, Cameron, Nikandish, & Shaveisi, 2016), which have explored similar results about graphs associated with finite groups

The properties of the independent set polytope, denoted as $\Gamma_{IP}(G)$, are being discussed.In this section, we present a comprehensive overview of the fundamental characteristics of the function.$\Gamma_{IP}(G)$.

Theorem 2.1: Consider a finite group G. The unicyclic property of $\Gamma_{IP}(G)$ holds if and only if G is isomorphic to either $\mathbb{Z}_3$ or S_3, where S_3 represents the symmetric group on three letters.

Proof. It is evident that the group $\Gamma_{IP}(\mathbb{Z}_3)$ can be represented as a cycle with a length of 3. The group $\Gamma_{IP}(S_3)$exhibits a single cycle of length 3 that is generated by the identity element, along with two elements of order 3. On the other hand, let us assume that the induced subgraph $\Gamma_{IP}(G)$ is unicyclic. Subsequently, we shall demonstrate the following:

1. The group$|G|$ does not possess any prime divisor p where p is greater than or equal to 5, as the graph $\Gamma_{IP}(G)$ exhibits unicyclic behaviour. Consequently, the cardinality of the group $|G| = 2^m 3^n$.
2. Let us consider a Sylow 3-subgroup, denoted as M, of the group G. Let us assume that the cardinality of set $|M|$ is greater than or equal to 9. If the group G contains an element x with an order of 9, then the subgroup generated by x, denoted as $\langle x \rangle$, will have at least two cycles in the permutation representation $\Gamma_{IP}(G)$. This leads to a contradiction. If the set M does not contain any elements with an order of 9, then every non-trivial element in M must have an order of 3. This in turn implies that the group $\Gamma_{IP}(G)$ must have at least two cycles, which contradicts the previous statement.
3. According to condition 2), it can be deduced that the order of the group G, denoted as $||G| = 2^m 3$. Additionally, it is known that G possesses a solitary Sylow 3-subgroup.
4. In a similar vein, it can be inferred that the set G does not contain any elements that have orders of both 4 and 6
5. Let $P = \langle x \rangle$ denote the distinct Sylow 3-subgroup of group G. The subgroup $P$is considered to be normal in the group G. According to the "'N/C' Theorem, it is

possible to embed the group $G/C_G(P)$ into the automorphism group of $\mathrm{Aut}(P)$. Let us assume the existence of an element $y$in the set$G \setminus P$
such that the elements x and y commute, i.e., $xy = yx$. The order of y is 2. Given that the orders of x and y are coprime, it follows that the order of xy is 6, which is deemed unattainable. This implies that the composition of the function $C_G(P) = P$. Furthermore, it is widely recognized that the automorphism group of the set $\mathrm{Aut}(P) \cong \mathbb{Z}_2$. Therefore, the cardinality of the quotient group $\left|\frac{G}{P}\right| = 1$either 2. In the aforementioned scenario, one. The condition $G = P = Z_3$ is satisfied as intended. In the latter scenario, it can be observed that the order of the group $|G| = 6$, and it can be readily demonstrated that $G = S_3$

Theorem 2.2 states that for a finite group G with order $n = p_1^{\beta_1} p_2^{\beta_2} \cdots p_m^{\beta_m}$, where $p_1, p_2, \ldots, p_m$and are distinct prime numbers and $\beta_1, \beta_2, \ldots, \beta_m$ are natural numbers, G possesses unique subgroups with orders $p_1, p_2, \ldots, p_m, u$ and $v$are non-adjacent in the graph $\Gamma_{IP}(G)$, where $u, v(\neq e) \in G \Leftrightarrow$ if and only if the greatest common divisor between the order of u and the order of v is one.

Proof. Assuming that the highest common factor between the order of element u and the order of element v is one, it follows that the intersection of the subgroups generated by u and$v$, denoted as ⟨u⟩ and ⟨v⟩ respectively, consists solely of the identity element e. Hence, the vertices u and v are not adjacent.

On the contrary, let us assume that u and $v$are not adjacent. If the greatest common divisor between the order of u and the order of $v$is not equal to one, then there exists a prime number p_i such that p_i divides the order of $o(u)$ and p_i divides the order of$o(v)$. This suggests that the elements $\langle u \rangle$ and $\langle v \rangle$ should possess a subgroup characterized by an order of p_i. Given that the group has a distinct subgroup with an order p_i, $|\langle u \rangle \cap \langle v \rangle| \geq p_i$, it follows that the intersection of the subgroups generated by elements u and v, Hence, the adjacency of u to v leads to a contradiction. Therefore, the greatest common divisor between the order of u and the order of v is one.

Theorem 2.3. In the case of a finite group G with an order expressed as $p_1^{\beta_1} p_2^{\beta_2} \cdots p_m^{\beta_m}$, where $p_1, p_2, \ldots, p_m$ are distinct prime numbers and$\beta_1, \beta_2, \ldots, \beta_m$ are natural numbers, it can be observed that the graph $\Gamma_{IP}(G)$ is connected and has a diameter that is less than or equal to 4. This holds true when G contains only subgroups with orders corresponding to $p_1, p_2, \ldots, p_m$.

Proof. Given that p_i is a divisor of the order of the group $|G|$, for $i = 1,2, \ldots, m$, it follows that there exists an element $x_i \in G$ such that the order of x_i is precisely p_i, for $i = 1,2, \ldots, m$. Let x_i and $x_j$denote two distinct elements in the group G, each having an order of p_i and p_j, where $i \neq j$ and $1 \leq i, j \leq m$. Let $N_i = \langle x_i \rangle$ and $N_j = \langle x_j \rangle$ denote the subgroups of G. Based on our presumptions, N_i and N_j denote distinct subgroups characterized by their respective orders, p_i and p_j. It can be inferred that both N_i and $N_j$are normal subgroups of the group G. Furthermore, it can be observed thatN_i and N_j, being a normal subgroup of G, possesses the property that its order is equal to the product of $p_i p_j$.Given that N_i and $N_j$are cyclic subgroups, it follows that their product, $N_i N_j$ is also a cyclic subgroup. Consequently, there exists an element y in $N_i N_j$with an order equal to the product of the orders of N_i and N_j, denoted as

$p_i p_j$. Based on our assumptions, the intersection of the sequence $\langle x_i \rangle \cap \langle y \rangle = \langle x_i \rangle$ and $\langle x_j \rangle \cap \langle y \rangle = \langle x_j \rangle$.This suggests that the product of $x_i y x_j$ forms a path in the graph$\Gamma_{IP}(G)$. Let u, vdenote two elements in the group G. Next, we consider the existence of p_i and p_j, where i and j are specific integers within the range of 1 to m, inclusive. It is required that p_i divides the order of u, while p_j divides the order of v. It should be noted that the expression $u x_i y x_j v$ represents a path connecting vertices u and v within the graph $\Gamma_{IP}(G)$. Therefore, it follows that the induced subgraph $\Gamma_{IP}(G)$ is connected and has a diameter of at most 4i.e, $(\Gamma_{IP}(G)) \leq 4)$.

Theorem 2.4.

The complete graphs K_5 and $K_{3,3}$ are examples of non-planar graphs. The non-planarity of $\Gamma_{IP}(G)$ can be observed in the case of an abelian group G with an order of either 12 or 18.

Proof. Case 1: The group G is cyclic.

For $|G| = 12, G \cong \mathbb{Z}_{12}$. Because G includes four elements of order 12 and a unique subgroup of order 2, $K_5$is a subgraph of $\Gamma_{IP}(G)$. According to Theorem 2.4, 2.4, $\Gamma_{IP}(G)$ is non-planar.

For $|G| = 18, G \cong \mathbb{Z}_{18}$. Because G includes six components of order $18, K_6$ is a subgraph of $\Gamma_{IP}(G)$. According to Theorem 2.4, $\Gamma_{IP}(G)$ is non-planar.

Case 2: G is not a cyclic group.

For $|G| = 12, G \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3$. G has six elements of order 6, hence K_6 is a subgraph of$\Gamma_{IP}(G)$. According to Theorem 2.4, $\Gamma_{IP}(G)$ is non-planar.

For $G = 18, G \cong \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_3$. G has eight elements of order 6, hence K_8 is a subgraph of$\Gamma_{IP}(G)$.According to Theorem 2.4, $\Gamma_{IP}(G)$ is non-planar.

Lemma 2.6. For any two finite groups H_1 and $H_2, \Gamma_{IP}(H_1) \cong \Gamma_{IP}(H_2)$ if $H_1 \cong H_2$

Proof. Assume that $f: H_1 \to H_2$ is a group isomorphism. Let $a, b \in H_1$be such that a is next to b in$\Gamma_{IP}(H_1)$. Since $\langle a \rangle \cong \langle f(a) \rangle$, for each $a \in H_1, |\langle a \rangle \cap \langle b \rangle| = |\langle f(a) \rangle \cap \langle f(b) \rangle| \geq 1$.In H_2, $f(a)$ is next to $f(b)$. As a result, $\Gamma_{IP}(H_1) \cong \Gamma_{IP}(H_2)$.

Remark 2.7: The converse of Lemma 2.6 is false. Consider $(\mathbb{Z}_8, +_8)$ and the quaternion group Q_8 with order 8. It should be noted that $\mathbb{Z}_8$ is not isomorphic to Q_8, but rather $\Gamma_{IP}(\mathbb{Z}_8) \cong K_8 \cong \Gamma_{IP}(Q_8)$.

Theorem 2.8: $G, \Gamma_{IP}(G)$ is a tree $\Leftrightarrow G \cong D_2$ or D_4 for any dihedral group G.

Proof. If G is isomorphic to D_2 or D_4, then its itersection power graph is K_2 or $K_{1,3}$. As a result, $\Gamma_{IP}(G)$ is a tree.

Consider the case where $\Gamma_{IP}(G)$ is a tree. Assume there is a prime number $p \geq 5$ that is a divisor of $|G|$. Because G contains an element of order $p, \Gamma_{IP}(G)$ contains $K_p(p \geq 5)$ as a subgraph. This $\Gamma_{IP}(G)$ is not a tree, which is a contradiction. As a result, $|G| = 2^n 3^m$, where $n \geq 1$ and $m \geq 0$ are two integers.

Assume$|G| = 2^n 3^m$, where $n \geq 3$ and $m = 0$. Then, as a subgraph, $\Gamma_{IP}(G)$ contains $K_{2^{n-1}}$. This $\Gamma_{IP}(G)$ is not a tree, which is a contradiction.

Assume $|G| = 2^n 3^m$, where $n \geq 1$ and $m \geq 1$ are integers. This means that G has an element w of order 3. The subgraph induced by $\langle w \rangle$ now contains K_3. This $\Gamma_{IP}(G)$ has K_3 as a subgraph, which is a contradiction. This means that G can be isomorphic to either D_2 or D_4

In this section, several finite groups are categorized as as whose $\Gamma_{IP}(G)$has a maximum book thickness of two.

Theorem 3.1. $m \geq 4, bt(K_m) = \left\lceil \frac{m}{2} \right\rceil$.

Theorem 3.2. If $G \cong D_{2n}$ where $n = 1,2,3,4$, $bt(\Gamma_{IP}(G))$ is at most two for a dihedral group G.

Proof. Given that the intersection power graph of D_2, D_4 and D_6 contains at least one edge, it can be concluded that the book thickness for these graphs is at least one. Each subgraph in this context refers to a subset of a larger graph, specifically the one-page embeddable graph associated with a given integer n. Therefore, the thickness of the lines representing the graphs in the book is one. The inclusion of the subgraph K_4 within the intersection power graph of D_8 implies, according to Theorem 3.1, that the minimum book thickness for this graph is two.

Theorem 3.3, if G is a finite abelian group and isomorphic to the trivial group of order one, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2, \mathbb{Z}_3 \times \mathbb{Z}_3 \times \cdots \times \mathbb{Z}_3$ or $\mathbb{Z}_4$., then the number of $bt(\Gamma_{IP}(G))$ (G)) is at most two.

Proof. The book thickness of the intersection power graph of the trivial group of order one is zero, as the graph is a null graph. Given the intersection power graph of the Cartesian product $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2$ and $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \cdots \times \mathbb{Z}_3$, it can be observed that there exists at least one edge. Consequently, the book thickness for these graphs is determined to be no less than one. The isomorphism between the intersection power graph of $\mathbb{Z}_4$ and K_4 can be observed. According to Theorem 3.1, the graph in question has a book thickness of two. In the current section, the values of $\alpha_0(\Gamma_{IP}(G))$ are achieved.

Theorem 4.1. For a finite group G with order$p_1^{\beta_1} p_2^{\beta_2} \cdots p_m^{\beta_m}$, where $p_1, p_2, \dots, p_m$are distinct prime numbers and $\beta_1, \beta_2, \dots, \beta_m$ are positive integers, the inequality $\alpha_0(\Gamma_{IP}(G)) \geq m$ holds.

Proof. Given that each p_i is a divisor of $|G|, G$, it follows that G possesses elements a_i such that the order of each element a_i, denoted as $o(a_i) = p_i$, for $1 \leq i \leq m$, where *i* ranges from 1 to *m*. It is important to highlight that the intersection of the sequence $\langle a_i \rangle \cap \langle a_j \rangle = \{e\}$, for each $i \neq j$, for every i not equal to j. The set $\{a_1, a_2, \dots, a_m\}$ is an independent set of the induced subgraph $\Gamma_{IP}(G)$.. Therefore, the desired outcome can be inferred.

Theorem 4.2. For a finite group with order $p_1^{\beta_1} p_2^{\beta_2} \cdots p_m^{\beta_m}$, where $p_1, p_2, \dots, p_m$ are distinct prime numbers and $\beta_1, \beta_2, \dots, \beta_m$ are positive integers, it is true that $\alpha_0(\Gamma_{IP}(G)) = m \Leftrightarrow G$ if and only if.The group G possesses a distinct subgroup characterized by order of p, where*i* ranges from 1 to *m*.

Proof. Let us consider a scenario in which the group G possesses a solitary subgroup that has an order denoted as p_i, where $i = 1,2, \ldots, m$. If the independence number of the induced subgraph □$_0$(Γ□□(□)) > □,□, then the graph G contains an independent set □ with at least □ + 1 elements. According to Theorem 2.2, the orders of elements in set □ are mutually prime. Given that the group □ has precisely m distinct prime divisors, it is not possible to identify m+1 elements within □ whose orders are mutually prime. The inequality □$_0$(Γ□□(□)) ≥ □ holds. Furthermore, according to Theorem 4.1, it can be concluded that the lower bound of □$_0$(Γ□□(□)) ≥ □ is greater than or equal to m. Therefore, the value of □$_0$(Γ□□(□)) ≥ □ is equal to m.

On the other hand, let us assume that the value of □$_0$(Γ□□(□)) = □. According to Cauchy's Theorem, the set □ contains elements that have an order denoted by □$_□$, where □ = 1,2, … , □. Consider a set of elements □$_□$ ∈ □, where □(□$_□$) = □$_□$, where □ = 1,2, … , □.

Let us consider a group □ that possesses two distinct subgroups, each having an order denoted by □$_□$, where i represents a specific index. Let the elements □$_□$ ∈ □ be such that the order of each element, denoted as □(□$_□$) = □$_□$. It is evident that the intersection of the sequences ⟨□$_□$⟩ ∩ ⟨□$_□$⟩ = {□} and so {□$_1$, □$_2$, … , □$_□$, □$_□$}is the independent set in in Γ□□(□) with □ + 1 elements. This observation leads to a contradiction. Hence, the group □ possesses a distinct subgroup of size □$_□$, where □ = 1,2, … , □

4. Result and Discussion

In this section, the results and interoperations of the analysis are documented. Said differently, by considering the sample universities results and graphical analysis are documented separately for each university.

4.1. Air University

Air University is a degree-awarding institute in Pakistan in Network Visualization. There are 408 items,20 clusters, and links connected are 2255, and the total link strength are 5923.An analysis was conducted on several authors' publication records, considering the number of documents they authored and the "Total link strength" metric. Among the authors, I.M. Qureshi emerged with the most documents, having authored 131 publications. Additionally, Qureshi exhibited a significant total link strength of 311, indicating strong connections and collaborations with other researchers in the network. Following Qureshi, A. Jalal ranked second with 90 documents and a cumulative link strength of 227. A.R. Javed secured the third position with documents and a cumulative link strength 192. Similarly, M.S. Arif had 53 documents and a total cumulative link strength of 150, while M.Y. Malik produced 58 documents with a total cumulative link strength 149. K.U. Rehman contributed 69 documents with a total link strength 147, while N. Naseer authored 58 documents with a total link strength of 133. Y.Y. Ghadi, with 27 documents, showed a high total link strength of 124, indicating strong collaborations despite a comparatively lower publication count. M.A. Khan and A. Raza published 46 and 35 documents, respectively, and shared a total link strength of 124.These findings shed light on these authors' research productivity and collaborative networks, emphasizing the significance of the number of documents authored and the strength of their connections with other researchers in the field.

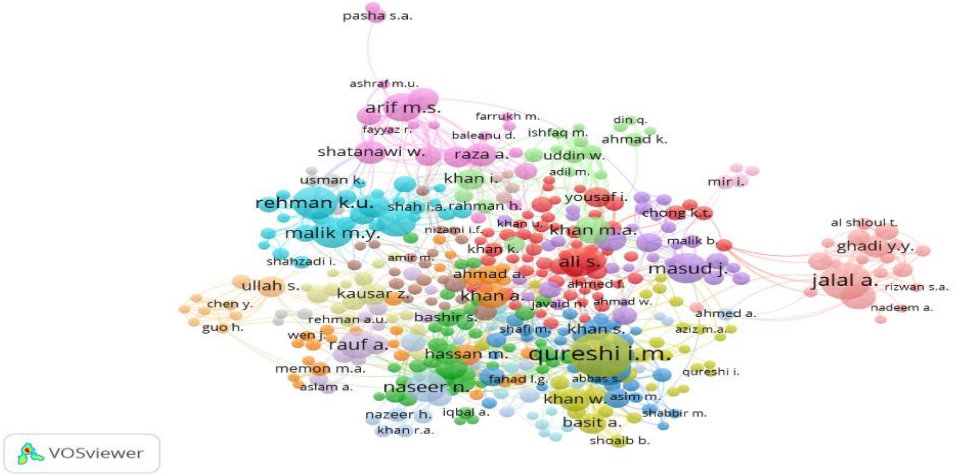


Figure 1: Co-Authorship Graph of Air University

Table 1: Collaborative Cluster Based on Productivity of Air University

High-Productivity		Moderate-Productivity	Low-Productivity		
Authors	Documents, Total Link Strength	Authors	Documents, Total Link Strength	Authors	Documents, Total Link Strength
Qureshi I.M	131	Arif M.S	53	GhadiY.Y	27
Jalal A.	90	Malik M.Y	58	Raza A	35
Javed A.R	72	Naseer N	58		
Rehman K.U.	69	Khan M.A	46		

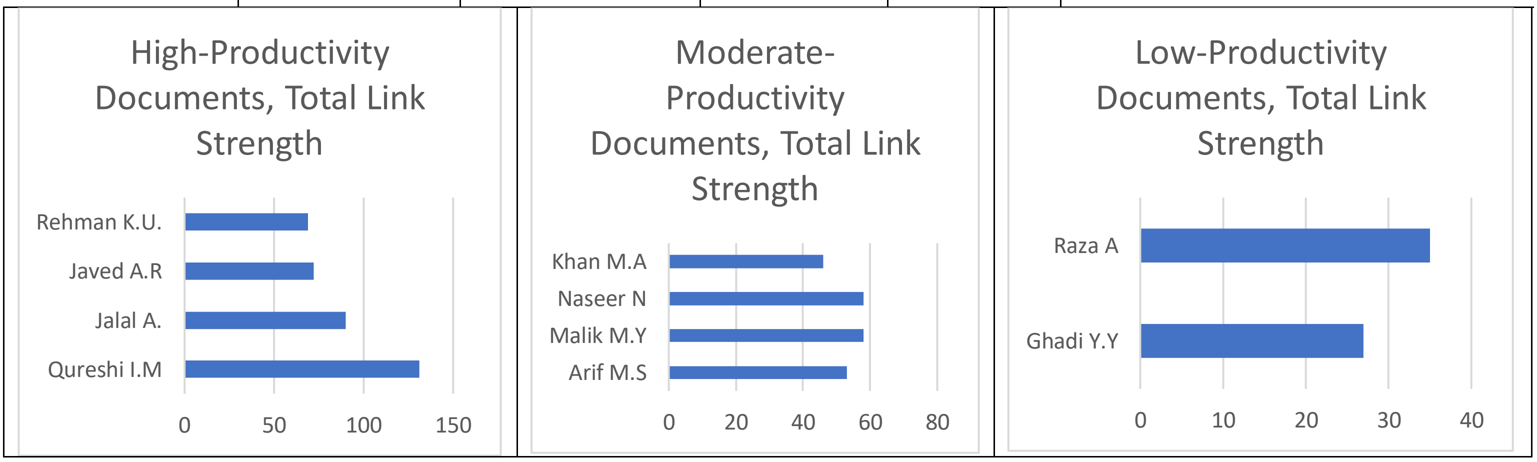


Figure 2: Collaborative Cluster Based on Productivity Graph of Air University

We can analyze the information based on the provided authors' data, documents, and total link strength to identify potential clusters or groups of authors. Clustering can help reveal patterns or associations among the authors based on their publication records and collaborative connections. Here is a clustering based on the provided data:

4.2. Bahria University

Bahria University consists of 404 items which make 22 clusters of the researcher's links which connect them are 2191 and the total link strength is 5010.An analysis was conducted on the publication records of several authors, taking into account the quantity of documents they authored and the "Total link strength" metric. Among the authors, M. Ramzan stood out with the highest number of documents, having authored 115 publications. Furthermore, Ramzan displayed a notable total link strength of an unspecified value, indicating strong connections and collaborations with other researchers in the network. Following Ramzan, K.N. Qureshi secured the second position with 73 documents and a total link strength of 169. Similarly, M. Usman ranked third with 52 documents and a total link strength of 167. G. Jeon contributed 58 documents and exhibited a substantial link strength of 158, while A. Ali authored 51 documents with a total cumulative link strength of 145.R.U. Haq produced 56 documents with a total link strength of 144, while A. Waqar had 42 documents and a total link strength of 122. S. Kadry authored 31 documents and exhibited a total link strength of 115. M. Hamid contributed 36 documents with a total cumulative link strength of 113, while A. Ahmad had 42 documents and a cumulative link strength of 100.S. Khalid authored 52 documents with a total link strength of 97, and S. Iqbal produced 42 documents with a total link strength of 95. J.D. Chung displayed 26 documents and a 93 total link strength. A. Shafee and M. Hussain published 26 and 39 documents, respectively, with total link strengths of 90 and 82. Additionally, S. Ahmad authored 29 documents with a cumulative link strength of 79, Y.-M. Chu had 20 documents with a 77 total link strength, and D.I. Godil contributed 27 documents with a 77 total cumulative link strength. These findings provide insights into these authors' research productivity and collaborative networks, highlighting the significance of the number of documents authored and the strength of their connections with other researchers in the field.

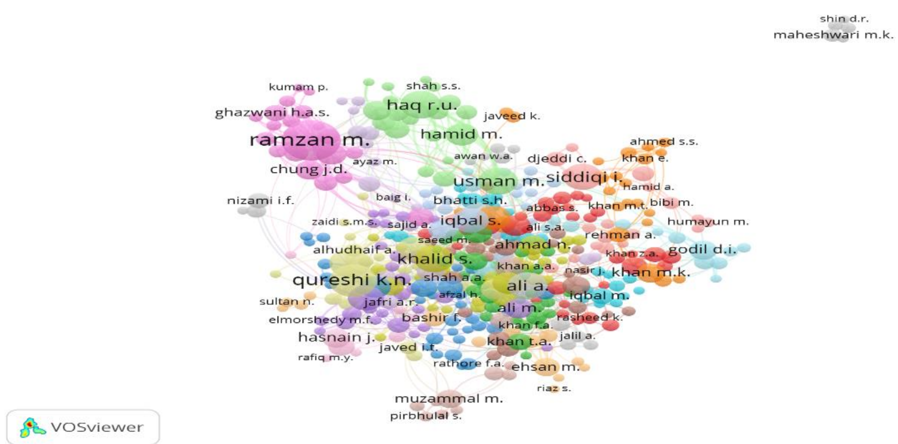


Figure 3: Co-Authorship Graph of Bahria University

Based on the given data of authors, documents, and total link strength, we can analyze the information to identify potential clusters or groups of authors. Clustering can help reveal patterns or associations among the authors based on their publication records and collaborative connections. Here is a clustering based on the provided data:

Table 2: Collaborative Cluster Based on Productivity of Bahria University

High-Productivity		Moderate-Productivity		Low-Productivity	
Authors	Documents, Total Link Strength	Authors	Documents, Total Link Strength	Authors	Documents, Total Link Strength
Ramzan M	115	Waqar A	42	Chung J.D	26
Qureshi K.N	73	Kadry S	31	Shafee A	26
Usman M	52	Hamid M	36	Hussain M	39
Jeon G	58	Ahmad A	42	Ahmad S	29
Ali A	51	Khalid S	52	Chu Y.-M	20

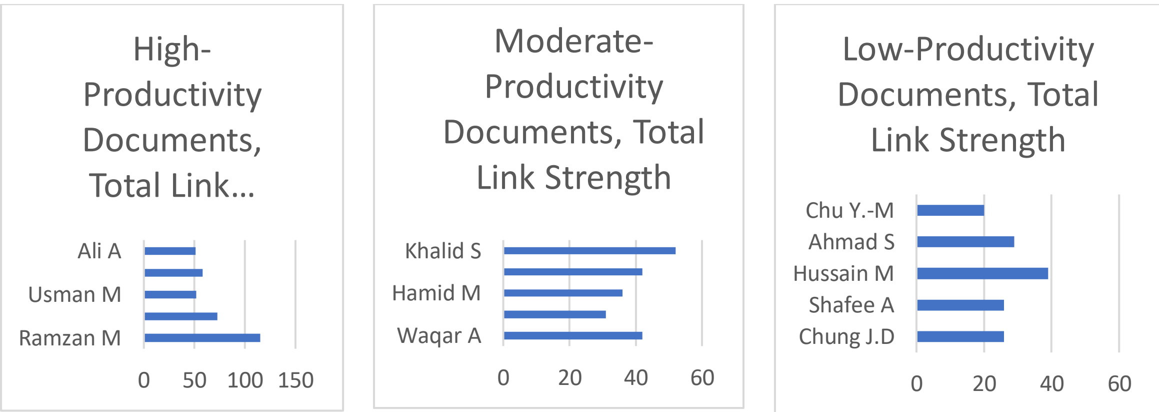


Figure 4: Collaborative Cluster Based on Productivity Graph of Beheria University

4.3. Comsats University Islamabad

Comsats University consists of 567 items, making 27 clusters of the researcher's links, which connect them with4191 and the total link strength is 1010. An analysis was conducted on the publication records of several authors, taking into account the amount of documents they authored and the "Total shared link" metric. Among the authors, K. Ayub emerged with the highest number of documents, having authored 59 publications. Additionally, Ayub displayed a total shared link value of 205, indicating strong connections and collaborations with other researchers in the network. Following Ayub, M.I. Khan ranked second with 57 documents and a total shared link value of 203. Similarly, S.U. Khan secured the third position with 55 documents and a total shared link value of 186. M. Imran contributed 48 documents with a total shared link value of 181, while A. Ali has authored 43 documents, with a collective shared link value of 152.A. Ahmad produced 52 documents with a total shared link value of 143, while T. Mahmood had 34 documents and a total shared link value of 131. A. Bokhari contributed 29 documents with a total shared link value of 120. J. Iqbal exhibited a total shared link value of 117, and S. Ahmad authored 41 documents with a total shared link value of 113.

Furthermore, M.A. Khan and A. Hussain shared a total link value of 112 and 107, respectively. S. Khan authored 39 documents with a total shared link value of 104, while M.A. Gilani contributed 25 documents with a total of 103.

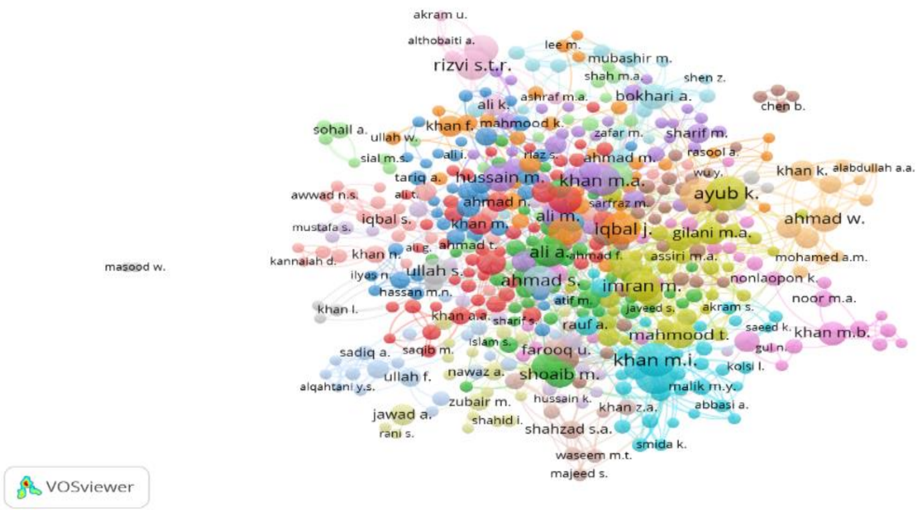


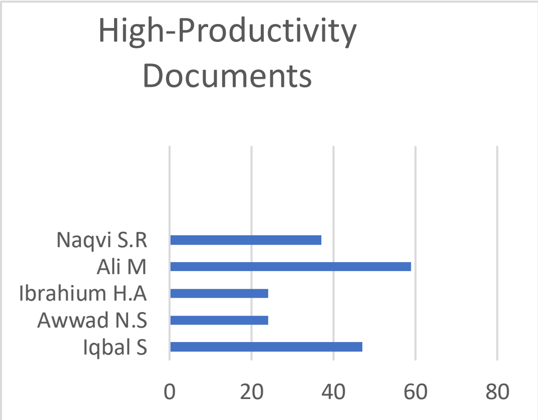
Figure 5: Co-Authorship Graph of COMSATS University, Islamabad

Based on the provided data of authors, documents, and total shared links, we can analyze the information to identify potential clusters or groups of authors. Clustering can help reveal patterns or associations among the authors based on their publication records and collaborative connections. Here is a clustering based on the provided data:

Table 3: Collaborative Cluster Based on Productivity of Comsat University

High-Productivity		Moderate-Productivity		Low-Productivity	
Authors	Documents / Total Link Strength	Authors	Documents / Total Link Strength	Authors	Documents, Total Link Strength
Ayub K	59	Iqbal J	49	0	0

Khan M.I.	57	Ahmad S	41	0	0
Khan S.U	55	Khan M.A.	39	0	0
Imran M	48	Hussain A	32	0	0
Ali A.	43	Khan S	39	0	0
Ahmad A	52	Gilani M.A.	25	0	0
Mahmood T	34				
Bokhari A	29				



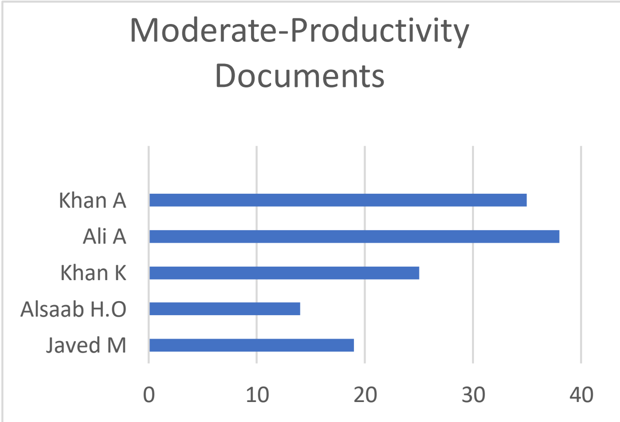


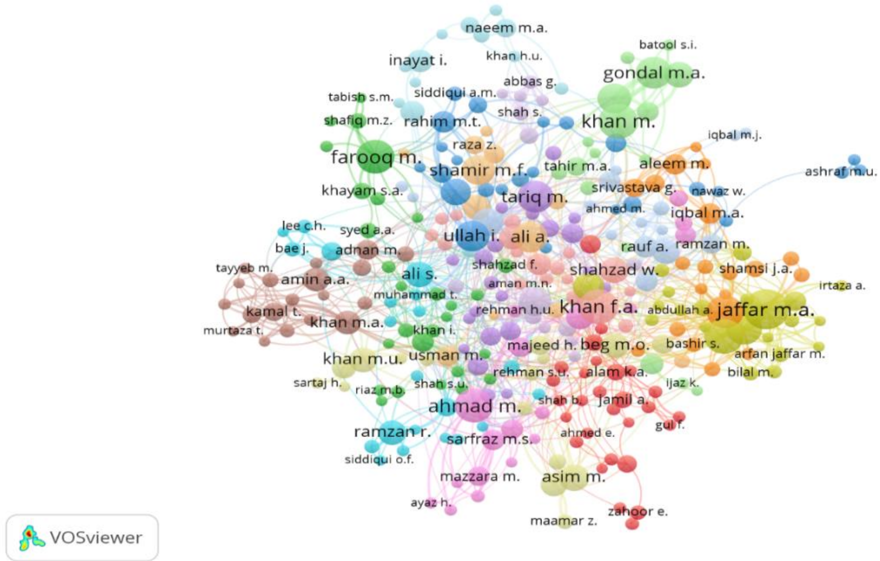
Figure 6: Collaborative Cluster Based on Productivity Graph of COMSAT University

3.4. Fast University, Islamabad

An analysis was conducted on several authors' publication records, considering the total number of documents they authored and the "Total link strength" metric. Among the authors, M.A. Jaffar stood out with the highest total number of documents, having authored 75 publications. Additionally, Jaffar exhibited an impressive total link strength of 1800, indicating strong connections and collaborations with other researchers in the network. Following Jaffar, A.M. Mirza achieved the second position, having authored 59 documents and possessing a total link strength of 143. Similarly, A. Hussain secured the third position with 49 documents and a total link strength of 135. M. Ahmad contributed 54 documents with a total link strength of 121, while M.A. Gondal authored 49 documents with a total link strength of 98. I. Hussain produced 42 documents with a total link strength of 96, while F.A. Khan and M.A. Khan both had 59 and 27 documents, respectively, with total link strengths of 93 and 89. M. Ali contributed 37 documents with a total link strength of 88, and M. Khan authored 57 documents with a total link strength of 81. Additionally, T. Shah had 33 documents with a total link strength of 79, while A.R. Baig and I. Ullah both exhibited a total link strength of 75 with 52 and 45 documents, respectively. S. Anwar authored 18 documents with a total link strength of 73, while S. Ali contributed 33 documents with a total link strength of 72. A. Ali and M. Asim both published 46 and 34 documents, respectively, with total link strengths of 71.

Furthermore, M. Farooq and M. Tariq both shared a total link strength of 71 with 54 and 48 documents, respectively. S. Khan and M. Khan authored 37 and 37 documents, respectively, with total link strengths of 69 and 70.

Based on the provided data of authors, documents, and total link strength, we can analyze the information to identify potential clusters or groups of authors. Clustering can help reveal patterns or associations among the authors based on their publication records and



collaborative connections. Here is a clustering based on the provided data:

Figure 7: Co-Authorship Graph of Fast University, Islamabad

Table 4: Collaborative Cluster Based on Productivity of Fast University

High-Productivity		Moderate-Productivity		Low-Productivity	
Authors	Documents / Total Link Strength	Authors	Documents / Total Link Strength	Authors	Documents / Total Link Strength
Jaffar M. A	75	Hussain I	42	Shah T	33
Mirza A.M	59	Khan F. A	59	Baig A. R	22
Hussain A	49	Khan M. A	57	Ullah I	25
Ahmad M	54	Ali M	37	Anwar S	18
Gondal M. A	49	Khan M	57	Ali S	23

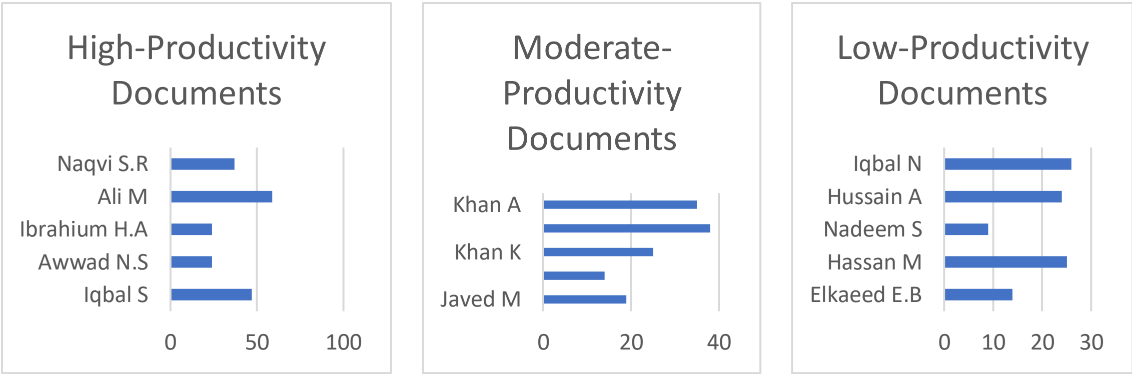


Figure 8: Collaborative Cluster Based on Productivity Graph of FAST University

3.5. International Islamic University, Islamabad

An analysis was conducted on the publication records of several authors, considering both the documents they authored and the "Total link strength" metric. Among the authors, T. Mahmood stood out with the maximum number of documents, having authored 128 publications. Additionally, Mahmood exhibited an impressive total link strength of 290I, indicating strong connections and collaborations with other researchers in the network. Following Mahmood, Z. Ali secured the second position, having authored 98 documents and possessing a total link strength of 219. Similarly, M. Arshad secured the third position with 75 documents and cumulative link strength of 208. I. Ahmad contributed 91 documents with a total link strength of 195, while N. Ali authored 75 documents with a total link strength of 180.K. Ullah produced 57 documents with a strength of 166, and A. Hussain had 64 documents with a total A. Ali contributed 42 documents with a total link strength of 131, while S. Khan exhibited a total link strength of 129. R. Ellahi authored 63 documents with a total link strength of 122, and A. Zeeshan contributed 55 documents with a total link strength of 113.Additionally, M. Ahmad had 40 documents with a total link strength of 111, while B. Uzair had 24 documents and a cumulative total link strength of 98. M.M. Bhatti contributed 35 documents with a total link strength of 90, and A. Ghani had 33 documents with a cumulative total link strength of 83. M. Ali and M.S. Khan shared total link strengths of 82, with 33 and 26 documents, respectively.Furthermore, A. Irshad authored 30 documents with a total link strength of 81, while A. Khan and A.A. Khan contributed 27 and 39 documents, respectively, with cumulative link strengths of 81 and 78.

Copyrights @Muk Publications

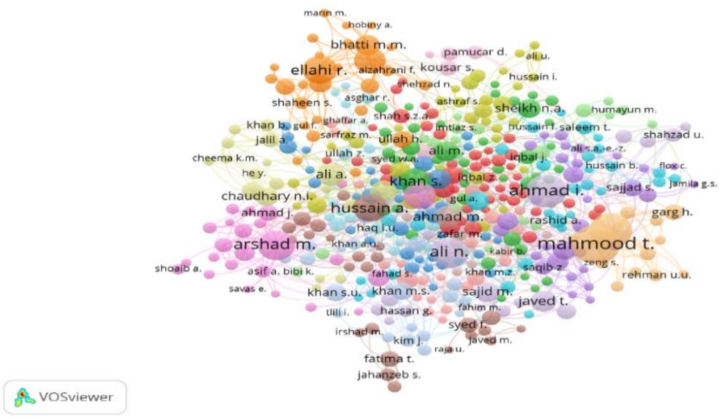


Figure 9: Co-Authorship Graph of International Islamic University

Based on the provided authors' data, documents, and total link strength, we can analyze the information to identify potential clusters or groups of authors. Clustering can help reveal patterns or associations among the authors based on their publication records and collaborative connections. Here is a clustering based on the provided data:

Table 5: Collaborative Cluster Based on Productivity of International Islamic University					
High-Productivity		Moderate-Productivity		Low-Productivity	
Authors	Documents / Total Link Strength	Authors	Documents / Total Link Strength	Authors	Documents/ Total Link Strength
Mahmood T	128	Ali A.	42	Uzair B	24
Ali Z	98	Khan S	40	Bhatti M.M	35
Arshad M	75	Ellahi R	63	Ghani A	33
Ahmad I	91	Zeeshan A	55	Ali M	33
Ali N	75	Ahmad M	40	Irshad A	30

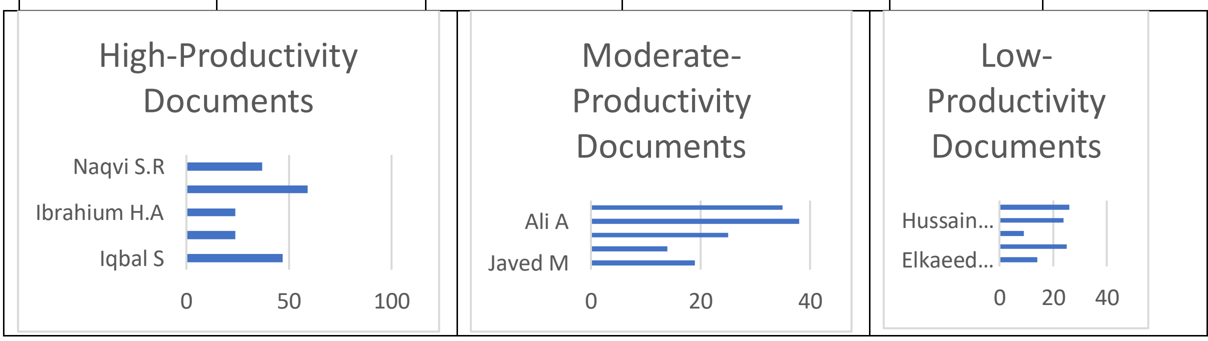


Figure 10: Collaborative Cluster Based on Productivity Graph of International Islamic University

3.6. National University of Science and Technology Islamabad

An analysis was conducted on the publication records of several authors, considering both the higher number of documents and a greater total link strength metric. Among the authors, S. Iqbal stood out with the most documents, having authored 47 publications. Additionally, Iqbal exhibited a substantial total link strength of 217, indicating strong connections and collaborations with other researchers in the network. Following Iqbal, N.S. Awwad and H.A. Ibrahim ranked second with 24 documents each and a total link strength of 162. Similarly, M. Javed contributed 19 documents with a total link strength of 131, while M. Ali authored 59 documents with a total link strength of 123. E.B. Elkaeed and H.O. Alsaab shared a total link strength of 105 and 97, respectively, with 14 documents each. S.R. Naqvi contributed 37 documents with a total link strength of 94, while Z. Said had 60 documents with a total link strength of 93. M.A. Khan and A.H. Khoja shared a complete link strength 90, with 33 and 27 documents, respectively. K. Khan authored 25 documents with a complete link strength of 86. Additionally, A. Ali contributed 38 documents with a link strength of 84, while A. Khan had 35 documents with a complete link strength of 80. M. Imran authored 40 documents with a complete link strength of 76, and M. Hassan contributed 25 documents with a total link strength of 73. M. Iqbal had 34 documents with a whole link strength of 72, while S. Nadeem authored 9 documents with a complete link strength of 69. Moreover, A. Hussain contributed 24 documents with a total link strength of 68, and N. Iqbal had 26 documents with a total link strength of 67.

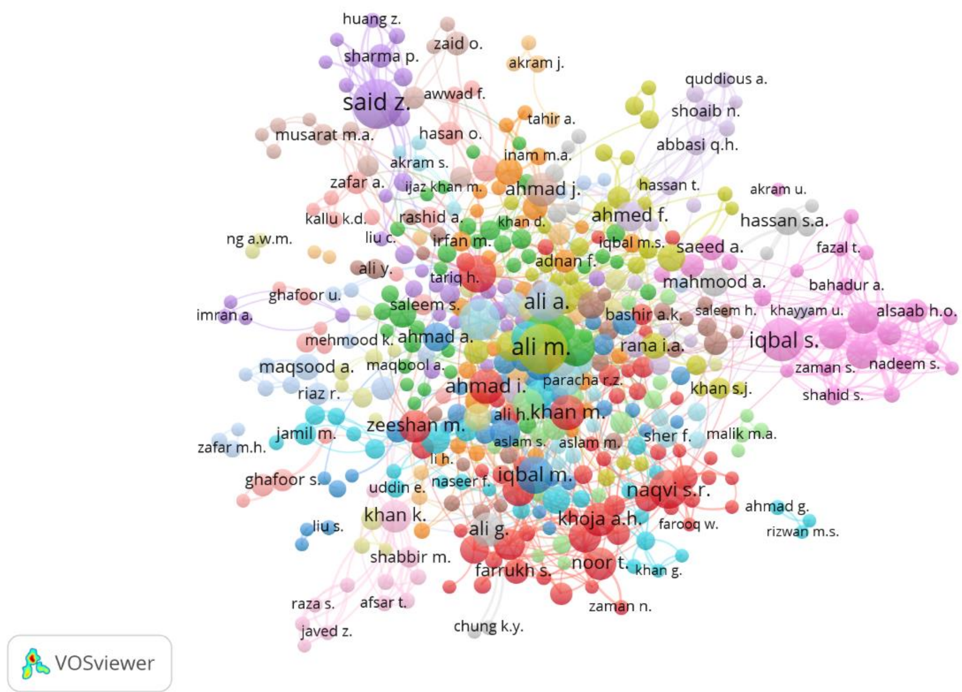


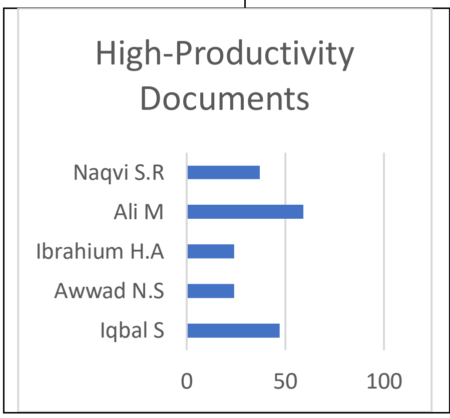
Figure 11: Co-Authorship Graph of NUST

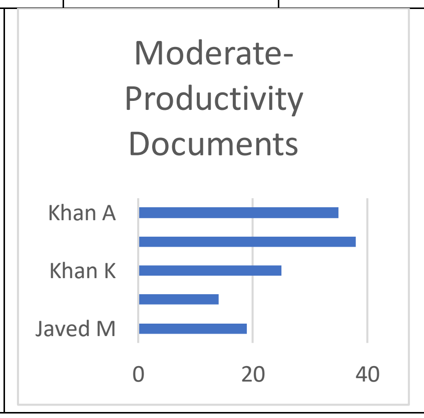
Figure 13: Collaborative Cluster Based on Productivity Graph of NUST University

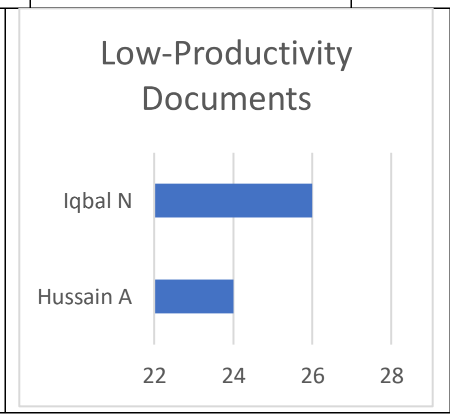
Based on the provided data of authors, documents, and aggregate link strength, we can analyze

Table 6: Collaborative Cluster Based on Productivity of NUST University

High-Productivity		Moderate-Productivity		Low-Productivity	
Authors	Documents / Total Link Strength	Authors	Documents / Total Link Strength	Authors	Documents/ Total Link Strength
Iqbal S	47	Javed M	19	ElkaeedE. B	14
Awwad N. S	24	Alsaab H. O	14	Hassan M	25
IbrahimH. A	24	Khan K	25	Nadeem S	9
Ali M	59	Ali A	38	Hussain A	24
Naqvi S. R	37	Khan A	35	Iqbal N	26





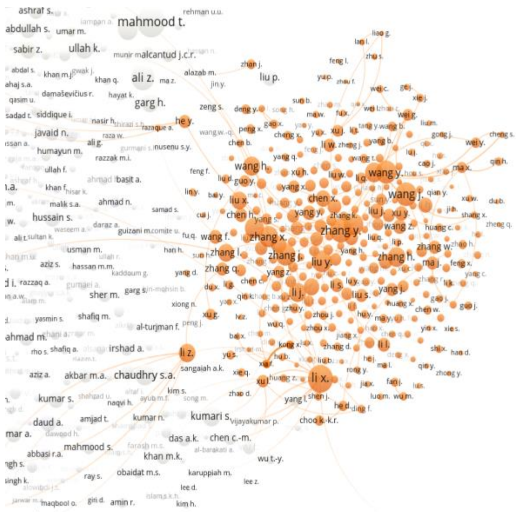


the information to identify potential clusters or groups of authors. Clustering can help reveal patterns or associations among the authors based on their publication records and collaborative connections. Here is a clustering based on the provided data:

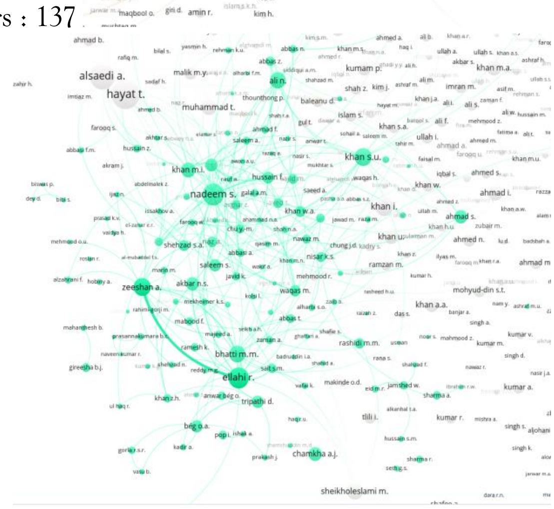
IIUI POWER GRAPH

An analysis was conducted on the publication records of several authors, considering both the higher number of documents and a greater total link strength metric. Total number of authors 999 link which is used to connect them are 13221 and total clusters are 17. In First cluster there are 317 members. In cluster 3 there are 136 members. in cluster 4 there are 92 members. In last cluster there are 67 members.

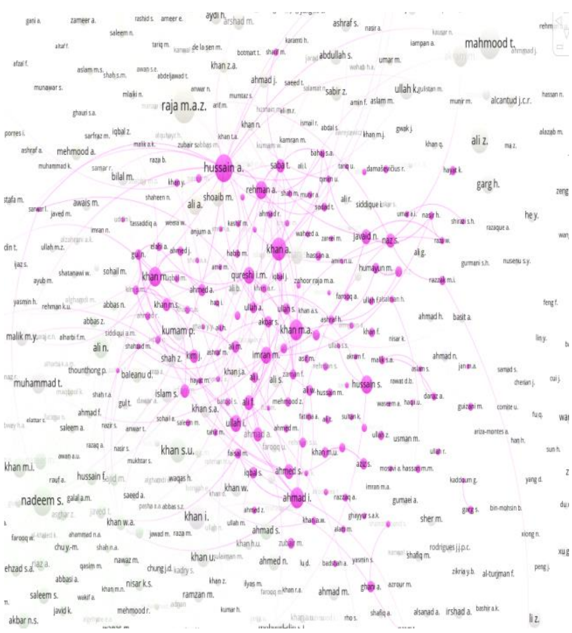
Cluster 1
Members : 317Cluster 2



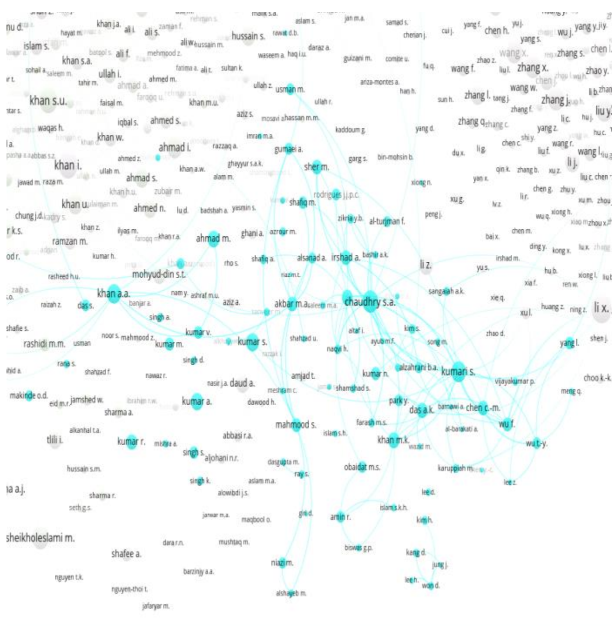
Members : 137



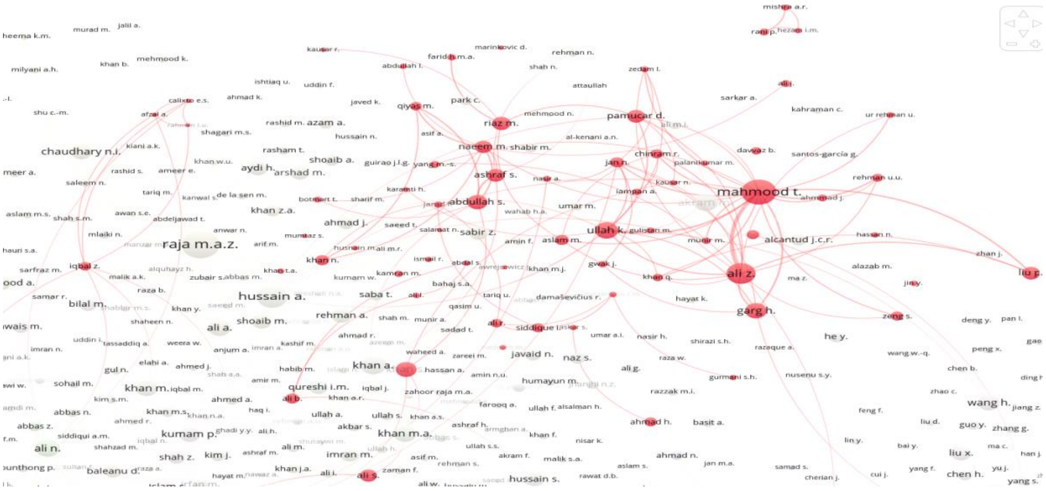
Cluster 3
Members 136



Cluster 4
Members 92



Cluster 5
Members 67

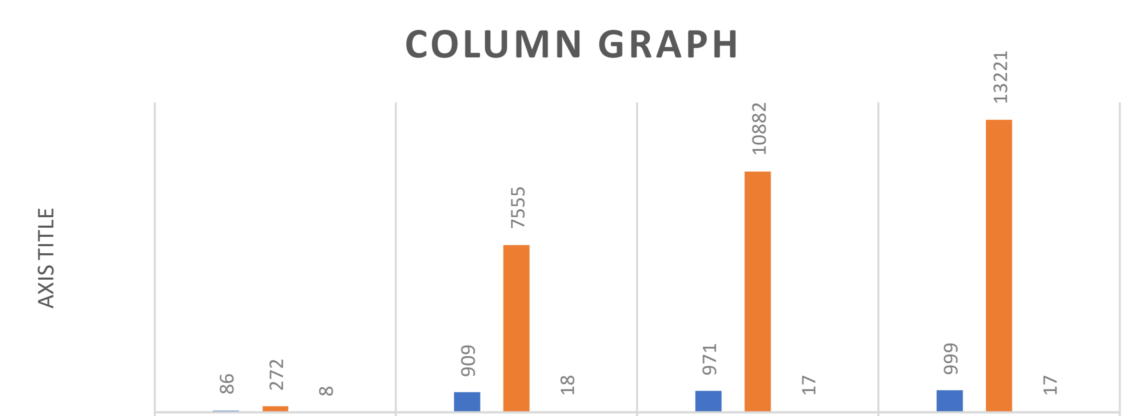


UNIVERSITIES BAR GRAPH

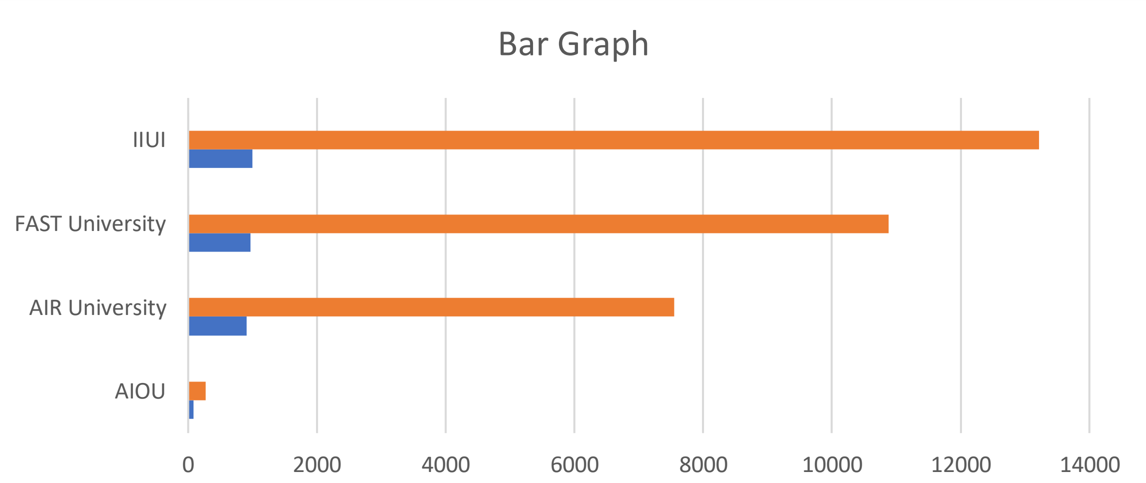
The utilization of charts facilitates the expeditious comprehension of data distribution and patterns, enables comparisons across various groups, and aids in the identification of potential outliers or extreme values. Therefore, we draw bar, whisker and box plot.

A bar chart is a visual depiction of data utilizing rectangular bars of differing lengths, wherein each bar corresponds to a distinct category or group, and the vertical dimension of the bar signifies the magnitude of that category. Bar charts are a valuable tool for the purpose of comparing various data points. We also draw bar chart for core universities based on authors, links, and clusters.

Further, whisker and box plot, commonly referred to as a box plot, serves as a visual representation that provides a concise overview of the distribution of a given dataset. The data visualization exhibits the median, quartiles, and potential outliers. In whiskers of a box plot, we present the to the minimum and maximum values within a specified range, of core universities.

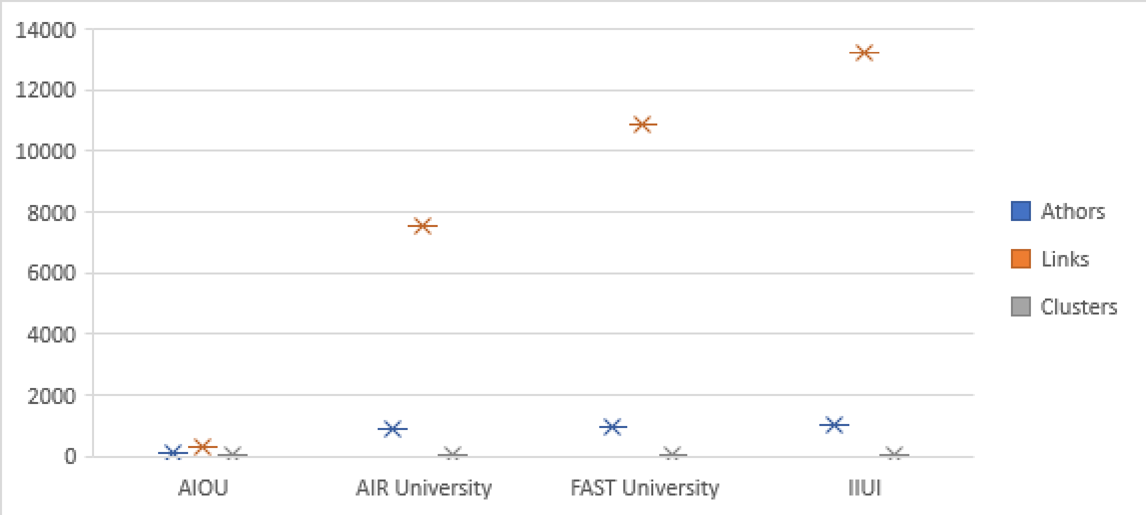


	AIOU	AIR University	FAST University	IIUI
Athors	86	909	971	999
Links	272	7555	10882	13221
Clusters	8	18	17	17



	AIOU	AIR University	FAST University	IIUI
Clusters	8	18	17	17
Links	272	7555	10882	13221
Athors	86	909	971	999

BOX AND Whisker Graph



Conclusions

This study introduced a unique approach for generating and analyzing scholar profiles using bibliometric data and information on their collaboration. The approach makes use of a variety of information that may be gleaned from an author's works and uses a new pipeline to group writers into categories, create profiles for each author, and calculate the ideal number of clusters automatically. We were able to group writers into subgroups, including the up-and-coming authors who stand out for their potential and dynamism, by using the provided technique on real dataset.

We crawled bibliographic data with time span from the Scopus digital library for the implementation of the recommended strategy and its assessment, and we fixed any discrepancies by using a method akin to collaborative filtering. Using graph representation and mining techniques, we were able to extract data on the rate of publications and emerging authors, the longevity of the top authors and their co-authorships over time, the influence a prosperous co-authoring community can have on someone's career, and other topics. We were also able to capture the social, individual, and time-related aspects of an author's impact.

The same process may be used to create other clusters from a different dataset, such as authors from a specific field, and it is anticipated that one of these clusters would include the rising stars cluster. Additionally, it may be used to identify and categorize groups of writers that have comparable productivity, influence, and sociability traits. In order to capture the temporal nature of a collaboration's success, new metrics characterizing an author emerged throughout this process, including time penalization in bibliometric, Power Graph, and social network analysis.

These features were used in yearly dataset construction, and using specific indicators, they were used to track each author's feature progression. These parameters made up the attributes on which we clustered the writers using the K-means technique. Using well-known clustering validity criteria and experimentation, the number of clusters was determined. It is our goal to use more sophisticated clustering algorithms (density-based, connectivity-based, probabilistic) that automatically determine the optimal number of clusters by using cluster coherence measures. By doing this, we will be able to get around K-means' drawbacks, such as the fact that it creates spherical clusters and requires the number of clusters as input. Depending on the feature properties of each cluster, cluster labelling was done. Furthermore, singular value decomposition was used in an effort to disclose the most important traits to the grouping. The findings indicated that the factor most important to an author's career is the number of articles penalized by time.

Our goals for the next work center on developing a classification mechanism that, given the necessary criteria, may assign an author to a certain group, i.e., embedding to the method the learning of a classifier based on the resultant clusters that will be seen as classes/categories. There are differences across disciplines including the natural, formal, social, and applied sciences, according to the study done on the dataset chosen by Pakistani affiliated academics. Since the suggested technique may be used with any group of writers, one of our future intentions is to concentrate on the researchers in a particular field. Additionally, it's crucial to use the whole crawled dataset and test out various grouping strategies. In order to achieve this, we intend to use a big data framework, such as Apache Spark(Panagopoulos et al., 2017), to attempt initially the use of all years' graphs and continuously an examination of the graphs' spectrum behavior over time. We also intended to use a spectral clustering approach, which was rendered impractical due to the volume of the graphs and the complexity of the eigen decomposition.

References

Aalipour, G., Akbari, S., Cameron, P. J., Nikandish, R., & Shaveisi, F. (2016). On the structure of the power graph and the enhanced power graph of a group. *arXiv preprint arXiv:1603.04337*.

Al-Hoorie, A. H., & Vitta, J. P. (2019). The seven sins of L2 research: A review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR Impact Factors. *Language Teaching Research, 23*(6), 727-744.

Bornmann, L., & Marx, W. (2015). Methods for the generation of normalized citation impact scores in bibliometrics: Which method best reflects the judgements of experts? *Journal of Informetrics, 9*(2), 408-418.

Campbell, D., Picard-Aitken, M., Côté, G., Caruso, J., Valentim, R., Edmonds, S., . . . Bastien, N. (2010). Bibliometrics as a performance measurement tool for research evaluation: The case of research funded by the National Cancer Institute of Canada. *American Journal of Evaluation, 31*(1), 66-83.

Chakrabarty, I., Ghosh, S., & Sen, M. (2009). *Undirected power graphs of semigroups.* Paper presented at the Semigroup Forum.

Cook, D. J., & Holder, L. B. (2006). *Mining graph data*: John Wiley & Sons.

Costas, R., van Leeuwen, T., & Bordons, M. (2010). Self-citations at the meso and individual levels: effects of different calculation methods. *Scientometrics, 82*(3), 517-537.

Cronin, B., & Overfelt, K. (1994). Citation-based auditing of academic performance. *Journal of the American Society for Information Science, 45*(2), 61-72.

Daud, A., Song, M., Hayat, M. K., Amjad, T., Abbasi, R. A., Dawood, H., & Ghani, A. (2020). Finding rising stars in bibliometric networks. *Scientometrics, 124*, 633-661.

Fitzgerald, H. E., Karen, B., Sonka, S. T., Furco, A., & Swanson, L. (2020). The centrality of engagement in higher education *Building the Field of Higher Education Engagement* (pp. 201-219): Routledge.

Fu, T. Z., Song, Q., & Chiu, D. M. (2014). The academic social network. *Scientometrics, 101*, 203-239.

Gibson, J., Anderson, D. L., & Tressler, J. (2014). Which journal rankings best explain academic salaries? Evidence from the University of California. *Economic Inquiry, 52*(4), 1322-1340.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. *Nature, 520*(7548), 429-431.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences, 102*(46), 16569-16572.

Kelarev, A. V., & Quinn, S. J. (2000). A combinatorial property and power graphs of groups. *Contributions to general algebra, 12*(58), 3-6.

Li, X.-L., Foo, C. S., Tew, K. L., & Ng, S.-K. (2009). *Searching for rising stars in bibliography networks.* Paper presented at the Database Systems for Advanced Applications: 14th International Conference, DASFAA 2009, Brisbane, Australia, April 21-23, 2009. Proceedings 14.

Long, P., Lee, T., & Jaffar, J. (1999). Benchmarking research performance in department of computer science, school of computing, national university of singapore: Technical Report.

Ma, Y., Yuan, Y., Wang, G., Bi, X., Wang, Z., & Wang, Y. (2020). Rising star evaluation based on extreme learning machine in geo-social networks. *Cognitive Computation, 12*, 296-308.

Marchionini, G., Plaisant, C., & Komlodi, A. (2003). The people in digital libraries: Multifaceted approaches to assessing needs and impact. *Digital library use: Social practice in design and evaluation*, 119-160.

McKenney, S., & Reeves, T. C. (2021). Educational design research: Portraying, conducting, and enhancing productive scholarship. *Medical Education, 55*(1), 82-92.

Panagopoulos, G., Tsatsaronis, G., & Varlamis, I. (2017). Detecting rising stars in dynamic collaborative networks. *Journal of Informetrics, 11*(1), 198-222.

Retzer, V., & Jurasinski, G. (2009). Towards objectivity in research evaluation using bibliometric indicators–A protocol for incorporating complexity (Vol. 10, pp. 393-400): Elsevier.

Royer, L., Reimann, M., Andreopoulos, B., & Schroeder, M. (2008). Unraveling protein networks with power graph analysis. *PLoS computational biology, 4*(7), e1000108.

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *Bmj, 314*(7079), 497.

Senanayake, U., Piraveenan, M., & Zomaya, A. (2015). The pagerank-index: Going beyond citation counts in quantifying scientific impact of researchers. *PloS one, 10*(8), e0134794.

Tagarelli, A., & Interdonato, R. (2015). Time-aware analysis and ranking of lurkers in social networks. *Social Network Analysis and Mining, 5*, 1-23.

Urooj, A., Khan, H. U., Iqbal, S., & Alghobiri, M. (2023). Exploring Author, Article, and Venue Feature Sets for Rising Star Prediction in Academic Network. *Journal of Scholarly Publishing*(aop), 1389-1414.

Waltman, L., & Van Eck, N. J. (2012). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology, 63*(2), 406-415.

Waltman, L., & van Eck, N. J. (2013). Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison. *Scientometrics, 96*, 699-716.